ESTABLISHING FORENSIC EVIDENCE VALUE THROUGH THE BAYESIAN FRAMEWORK AND THE LIKELIHOOD RATIO

Jovana Vujošević, MSc
Božidar Banović, PhD[1]
Faculty of Security Studies, University of Belgrade, Serbia

Abstract: Forensic science refers to the utilization of scientific methods in a lawful setting. Recent events have raised worries about applied techniques in forensic evidence analyses, specifically in which ways forensic scientific evidence is interpreted and presented in the court. Reports have recognized issues as the means of deficiency in numerous jurisdictions in regard to the amount of evidence requiring processing, insufficient standardization across laboratory facilities and professionals, and inquiries concerning the analysis, understanding and presentation of evidence. Questions have arisen about the main scientific foundation for forensic exam assessments on evidence types. Statistics has appeared as a crucial discipline for assisting the forensic community in addressing these difficulties. The essential standard components of statistical analysis represent study design, data collection and further analysis, statistical interpretation, outlining and reporting final results. This article explores the important role of forensic evidence, the diversity of forensic fields, current achievements and their limitations, and the expected commitments of more thorough statistical methods, particularly Bayesian approaches and the likelihood ratio in the analyses, interpretations, and forensic evidence report.

Keywords: Bayes factor, likelihood ratio, forensic evidence, DNA, probability, court.

INTRODUCTION

This article aims to explore the role of statisticians in forensic science reform and portray recent research efforts that have been applied. Statisticians have been involved in forensic and legal evidence since the mid-19th century. In Europe, both Francis Galton and Adolphe Quetelet put forth the idea

---

[1] banovicb@fb.bg.ac.rs

of uniqueness regarding the fingerprints. The contribution of statistics in forensics, for the most part, stayed limited until the usage of genetics and likelihood ratios in evaluating the strength of evidence - a possible match between DNA from a crime scene and a suspect (Pierson & Kafadar, 2016).

Various models have been suggested to evaluate and demonstrate the proposition value of forensic evidence. However, the statistical performance of inferences in forensic science includes a layer of difficulty that most scientific researchers and court officials are not equipped to handle (Neumann & Ausdemore, 2020).

For proper evidence assessment, you have to study probabilities of the evidence under two different propositions. These propositions are normally given by the prosecution and the defence. There are advanced statistical methods for doing this, which are constructed on likelihood ratios or Bayes factors. Estimations dependent on them may in some cases be forthright. However, there can appear some non-standard issues that need to be considered.

The likelihood ratio (LR) is utilized to assess the strength of the evidence and a reference specimen, whether they originate from common or different sources (Meuwly, Ramos & Haraksim, 2017).

Since each case represents a unique set of settings and findings, it is hard to think of a standard structure for assessment. The forensic science association has recognised the hierarchy of propositions (Evett, Jackson & Lambert, 2000). The higher the propositions are within the hierarchy, against which the researchers are capable to evaluate their outcomes, the more valuable the testimony will be to the court of law (Biedermann et al, 2016). We study a template for Bayesian Network construction that permits adequate adaptability to address most cases, but enough cohesion and structure that the information flow in the Bayesian Network is promptly perceived. We featured seven stages that can be utilized to construct a Bayesian Network inside this structure and validate how they can be applied, using a case example (Taylor, Biedermann., Hicks & Champod, 2018b).

IMPORTANCE OF STATISTICS IN FORENSIC SCIENCES

Whether forensic examinations in a specific area are reliable does not indicate whether they are valid or correct. If a forensic examiner determines that a latent fingerprint print at the crime scene comes from the same source as a test impression made by the suspect, we have to know how precise that conclusion is for making an educated judgment about the weight of evidence (Stern, 2017).

The question *"what to validate?"* centres around the validation methods and criteria, while the question *"how to validate?"* manages the implementation of the validation protocol.

After the validation protocol and report are proposed, they can be implemented to the forensics - creating and validating likelihood ratio methods for the assessment of the evidence strength at source level following these propositions (Meuwly et al., 2017):

H_1/Hp: The trace and reference originate from the same source;

H_2/Hd: The trace and reference originate from different sources.

The field of forensic science, similar to other areas involving human findings, has the potential to be compromised by selective perceptions and other factors which can impact how data is gathered, examined and findings are connected. Statisticians can monitor a group that study selective perceptions and a group that tries to establish what information should esteem important for a specific task (Stern, 2017).

At crime laboratories, quality maintenance of forensic examinations requires a quality affirmation program that incorporates reanalysis of evidence and verification of analyses conclusions. Statistical methods through designing a quality affirmation program can improve the proficiency of laboratory operations while simultaneously guaranteeing the precision of conclusions.

Several studies show a difficulty that court members can have in comprehending statistical ideas like the likelihood ratio and Bayes factor. Statisticians are creating ways to deal with introducing quantitative evaluations of evidence and within the design and analysis of courtroom, studies to assess the effectiveness of alternative approaches (Stern, 2017).

A final aim is to help forensic scientists in their quest of reliable and accurate analyses of forensic evidence.

BAYESIAN NETWORKS

If every single proposition is valid, the statistician allocates the probability of the forensic evidence to develop a likelihood ratio (LR).

This is performed in Bayesian Networks which are truly significant tools to pre-assess cases, because they are organized in an unprejudiced way, assessing every single possible result, consequently permitting the probabilities to be appointed without prior knowledge of results (Gill et al., 2020).

Bayesian Networks can also indicate how a lack of data and information could impact a particular case. Taylor, Kokshoorn & Biedermann (2018a) stated: “If there is a paucity of data used to assign a probability to which the LR is particularly sensitive, then this may indicate that the opinion of the scientists requires careful investigation of its robustness”. These issues ought to be accounted for before being heard in court.

The benefit of Bayesian Networks is that they can be utilized to assess complex outcomes that would be very difficult to assess by deriving formulae (Gill et al., 2020). A set of input probabilities are utilized to transmit an output that is established upon the propositions stating prosecution and defence views respectively. The output probabilities provide data about the probabilities of the forensic evidence results (e.g. quantity and profile with given characteristics) adapted on alternate activities (Taylor et al., 2018b).

Bayesian networks with a graphical approach of demonstrating and directing complex probability evaluations bring numerous ways how it can be constructed (Biedermann & Taroni, 2012). One concern of this is that there is a wide range of methods of probabilistically assessing the same set of findings. It ought to be noted, nonetheless, that different BN architectures may reflect different assumptions and underlying assessments, which may prompt contrasts in allocated values of evidence. Therefore, with further analyses, we can determine how the assessment of robust strength of evidence (in regards to a likelihood ratio) is to probability assignments underlying the nodes of the BN (Taylor, Hicks & Champod, 2016).

Upon facing new evidence, scientists calculate their weight of evidence as a personal likelihood ratio. Following Bayes rule, they will multiply their previous (prior) odds by their respective LR to obtain their restructured (posterior) odds reflecting their re-examined degrees of belief concerning the case being referred to (Lund & Iyer, 2017).

We are familiar with the work that focuses on the construction of Bayesian Networks in the law field.

There is, however, limited research on the integration of biological forensic results into a BN that reflects on propositions regarding activities (Gill et al., 2020).

STATISTICAL DNA ISSUES

The significance of the DNA evidence concerning an activity has to be studied according to highlights other than its profile, for example, the amount of DNA, level of degradation and its complexity (as numerous contributors). The outcomes of one or multiple DNA profiles must then be considered in light of the tendency for DNA to transfer and carry onto objects, through potentially multiple intermediaries. It is the probability of the results given each of the opposing actions – one would portray a certain component of the wrongdoing, and the other would identify with a means of innocent acquisition which is then delivered to the court (Gill et al., 2020).

It is necessary to assess evidence within a coherent structure that is provided by the 'hierarchy of propositions'. In that manner, the court will not estimate the value of evidence of the DNA profile regarding its source to the 'activity' that led to the DNA transfer (Buckleton et al., 2014).

At the point when forensic experts are evaluators, they assign a likelihood ratio that is characterized as the ratio of two probabilities:

- The probability of the results assumed that one proposition is true with the conditioning information;
- The probability of the results assumed that the other proposition is true with the conditioning information.

Clarifying the presence of DNA is more suited to the role of *investigators* – which happens before a defendant being put forward for prosecution. It is necessary to establish whether the results have higher probability within one particular activity level proposition over the other and in which quantity, to assist the court with addressing the case on which they are working on.

Concerning the case date, only the relevant case circumstances are vital for the *evaluation of results*: for instance, relevant information regarding the alleged activities, the timing, and the individual's things in the bag, what the suspect stated regarding the crime incident. A case of biased information, that is not required nor needed, would be that an eyewitness has identified the suspect (Gill et al., 2020).

There are many options for the use of verbal scale, but we have chosen the published ones by ENFSI (2015): neutral support (LR= 1); limited support (LR= 2–10); moderate support (LR= 10–100); moderately strong support (LR= 100–1000); strong support (LR= 1000–10000); very strong support (LR= 10.000–1.000.000) and extremely strong support (LR> 1.000.000). To avoid the misunderstanding and due to difficulties in describing the large numbers with words in a meaningful way, verbal scale finishes with the option of LR being greater than one million.

For instance: My LR is in the order of 70. In other words, the results are much more probable (in the order of 70 times) if the proposition that 'A' is true than if the alternative 'B' is true. Agreeing with ENFSI, these results indicate moderate support for the first proposition "A" over the alternative "B" (Gill et al., 2020).

Therefore, the higher the likelihood ratio, the stronger the evidence will be in support of the hypothesis that the source of the evidence-DNA sample and the suspect are the matching person (Weir, 2007).

Hp: the perpetrator left the bloodstain (views of prosecution in a criminal trial)

Hd: some other person left the bloodstain (views of defence in a criminal trial)

It will be quite useful if the prior odds are accessible on the hypothesis that the two DNA profiles derive from the same source. Prior odds represent the odds that the two pieces of DNA evidence derive from the same individual based on information other than the DNA. While, when the DNA information is incorporated in the analysis, we can calculate the posterior odds (NRC, 1996a).

$$\textit{Posterior odds} = \frac{\Pr(Hp|E)}{\Pr(Hd|E)}$$

Bayes' theorem:

$$\frac{\Pr(Hp|E)}{\Pr(Hd|E)} = \frac{\Pr(E|Hp)}{\Pr(E|Hd)} \times \frac{\Pr(Hp)}{\Pr(Hd)}$$

$$\text{POSTERIOR ODDS} = \text{LR} \times \text{PRIOR ODDS}$$

For calculating the probability, we should obtain prior probability and Bayes' theorem.

It is necessary to present the posterior probabilities corresponding to a prior range. A prior probability could be influenced by subjective evaluation of the examined evidence. Likewise, there is no need for non-DNA evidence to be presented first. It might be perplexing for a court to hear prior odds assigned by one professional, then receive information about a likelihood ratio from a different expert, continuing with more non-DNA evidence. Due to all this, we consider it best, if Bayes' theorem is used, to show posterior probabilities for a range of priors (NRC, 1996a).

This method of presentation is defined as "variable-prior-odds". During this process, statisticians do not apply their prior odds or focus on court members formed prior odds for substitution into Bayes' rule. The results are presented with a table or graph indicating how the posterior odds change as a function of the prior odds (NRC, 1996b).

While focusing on the question of uniqueness, we have determined two approaches – the first, as enquire for the probability that a given profile is unique and the second, a more complex that deals with the probability that no two profiles are identical. The first approach is expected to be inquired much more often in forensic involvement.

Due to the continuously increasing number of available loci for forensic DNA analysis, we can imagine an option of how the individual's profile could be unique (excluding identical twins and close family members). Assuming that, in a population of unrelated people (N), a particular DNA profile has its probability (P). Before a suspect is profiled, this probability for a *particular* profile, which has been observed within evidence, is not unique for most NP (NRC, 1996a).

BAYESIAN NETWORK CONSTRUCTION WITH FOCUS ON ACTIVITY LEVEL PROPORTIONS

We have chosen seven stages that depict the overall procedure of Bayesian Network construction while assessing forensic biological results taking into consideration of opposing propositions regarding activities:

1 - Outline the principle proposition node

2 - Outline activity node(s)

3 - Collect similar findings

4 - Outline findings node(s)

5 - Outline transfer and persistence node(s)

6 - Outline root nodes(s)

7 - Check for complete support within the Bayesian Network (Taylor et al, 2018b)

An example: A 20-year-old female (X), who lived with her biological mother (M) and father (F) stayed for a weekend at her elder brother's (Y) house. A girl phoned the police and stated that her sibling has harmed her in the genital region, over her clothes. Police arrived immediately. Her underwear was taken as evidence and a reference from the girl as well. The police then arrested the brother and took his DNA as a reference sample. The following information was gathered:

- The prosecution: The brother harmed her in the genital region, over her clothes.
- The defence: Girl was staying at the brother's home, but no harm was done.

1 - Outline the principle proposition node

- Y harmed X in the genital region, over her clothes.
- X was staying at the Y's home, but no harm was done.

2 - Outline activity node(s)

In this particular situation, the third person that we will need to consider is required because DNA evidence is Y-STR and Y is the son of F, therefore they share a Y-STR profile. As X generally lives with F, it might be that the male DNA discovered on the clothing of X comes from F (Fig. 1).

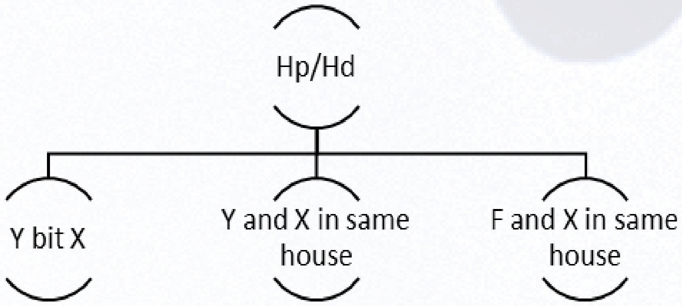


Figure 1: *Bayesian Network after the initial two steps of construction.*

3 - Collect similar findings

In our case, the findings were combined of the two tape lifts of the exterior surface of the underwear. In both cases, the results were that high levels of X's DNA were present. On one of the tape lifts, low levels of male DNA were detected and the subsequent Y-chromosome profiling generated a Y-STR profile.

4 - Outline findings node(s)

From this particular case, there are two results, the first is a positive RSID result for saliva which was gathered from the exterior surface of the crotch and underwear and the second is a low level of a male DNA found, which matched Y-STR (short tandem repeat) DNA profile with Y. Because Y and F are closely related, they are expected to have the same Y-STR profile which further defined the result as 'Family YSTR profile'.

5 - Outline transfer and persistence node(s)

All three activities can lead to the detection of the family Y-STR profile on the exterior surface of X's underwear and for better clarity, we could add a node to gather the two sources of Y's DNA before the YSTR findings node. Only the activity of biting would lead to the presence of Y's saliva on X's underwear and so only one path from this activity node to the RSID node can be presented.

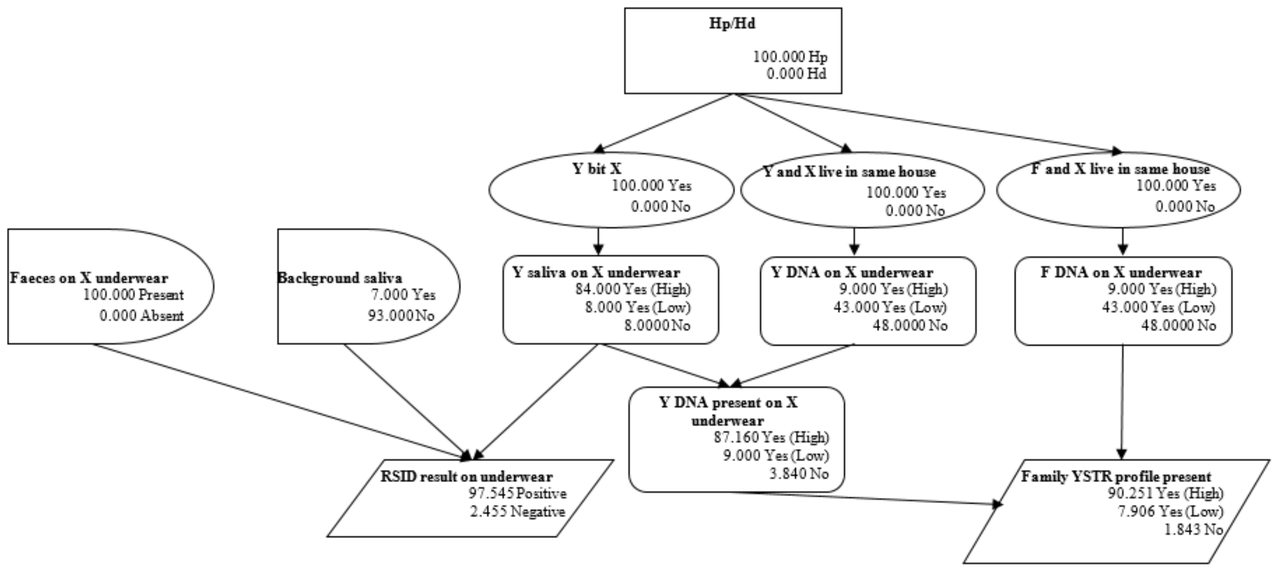
6 - Outline root nodes(s)

For instance, we could include in nodes the possibility of contamination, or coincidentally matching background DNA to the YSTR results node. Conversely taking into consideration that the alternate possibility of the DNA source is F, the rarity of contamination in this case or the profile frequency in contrast to most of the other transfer and persistence probabilities and knowing the fact that defence and prosecution are accepting the result to have arrived either Y (by innocent means or not) or F, the further nodes will contribute little to the BN and have an immaterial impact on the strength of the results (LR). We can include two root nodes, which will focus on the effect that the presence of faeces and general background levels of saliva on underwear can have within an analysis.

7 - Check for complete support within the Bayesian Network

All results ought to be recognisable under either proposition. This usually means that there will be two opposing activities that can proceed to the same finding.

From our BN construction, it is obvious that numerous routes indicate the presence of the family YSTR on the underpants of X, while some align with defence statement other support criminal activity. On one side, the RSID test node follows the route that confirms criminal activity, while on the other two sides there are two root nodes which represent two other routes (Fig. 2). This implies that the BN will not result at all probability being present on one of the states in the propositional node when findings nodes are instantiated.



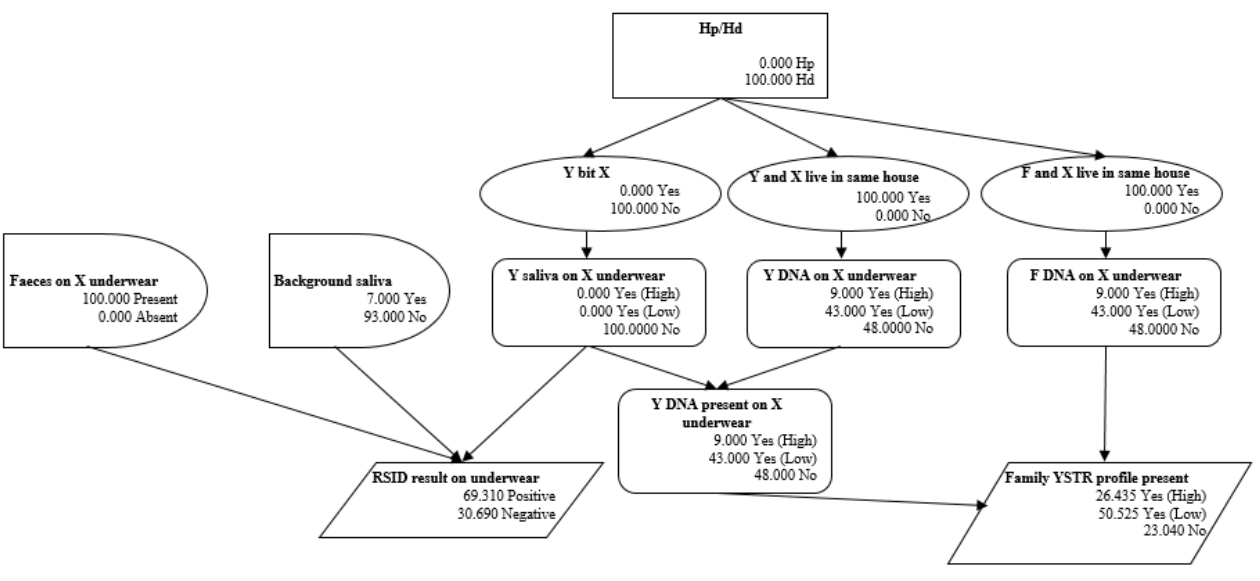


Figure 2: *BN with Hp (upper) and then Hd (lower) instantiated in proposition node - values rely upon the probabilities entered into conditional probability tables.*

$$\frac{\Pr(Hp|E,I)}{\Pr(Hd|E,I)} = \frac{\Pr(E|Hp,I)}{\Pr(E|Hd,I)} \times \frac{\Pr(Hp|I)}{\Pr(Hd|I)}$$

I illustrates case information and has been incorporated to underline its significance for activity level propositions.

When appointing probabilities to the propositional node it is often practised to arrange values equally (0.5 for both Hp and Hd). Therefore, the formula is simplified as (Taylor et al, 2018b):

$$\frac{\Pr(Hp|E,I)}{\Pr(Hd|E,I)} = \frac{\Pr(E|Hp,I)}{\Pr(E|Hd,I)}$$

as the ratio of the prior probabilities is one. A reminder that this is just with regards to the BN, as it is being applied by scientists and does not concern the probabilities doled out to the propositions by the court. Scientists look for the likelihood ratio, which is the principal term in the equation above. Considering the link between likelihood ratio and posterior odds, there are two ways of acquiring the numerical estimate of the likelihood ratio. One is to determine the ratio of the probabilities related to the findings in case (results node) when the proposition node is instantiated first in the Hp and then in the Hd.

In this manner, we gather probabilities of the evidence concerning each proposition and therefore one can estimate the likelihood ratio. BN construction with numerous results nodes may require an additional node with role to combine the findings (Fig. 3).

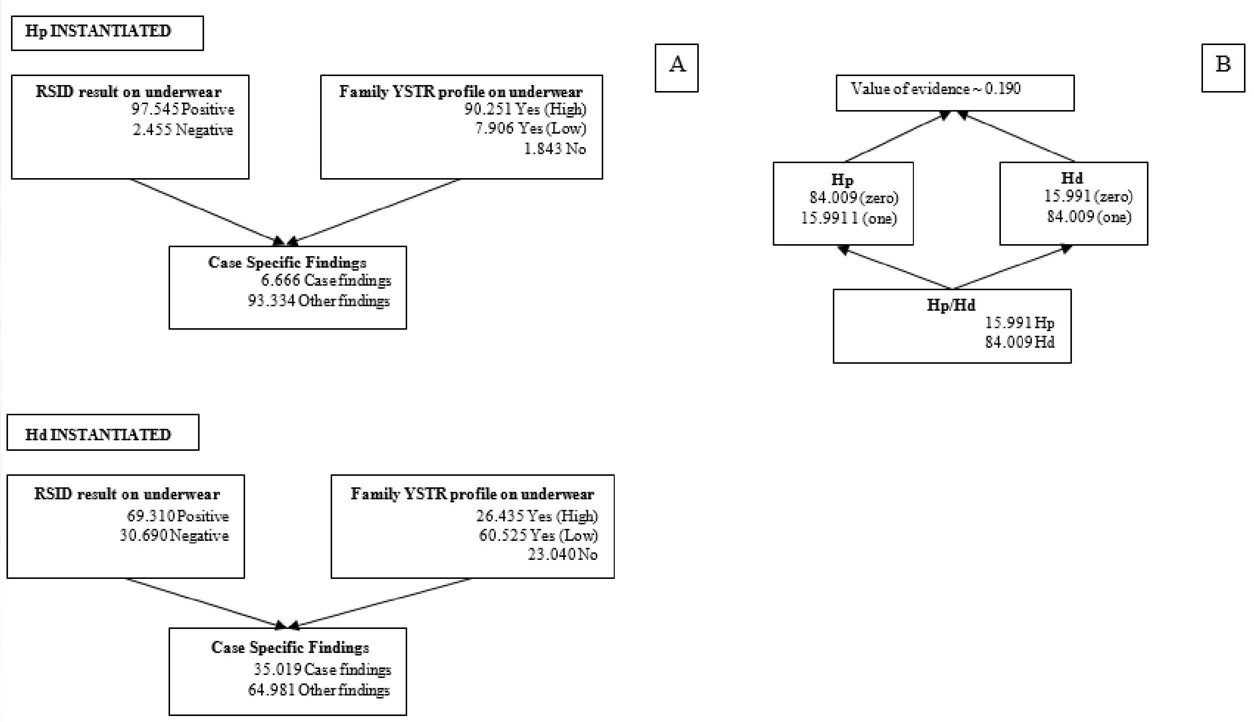


Figure 3: *Two available methods for assessing the strength of the evidence, using the BN from Fig 2*

a) Estimating LR (results node) - presenting the lower results nodes of the BN and the new specific results node;

b) Estimating the posterior odds (function node) – presenting the proposition node of the BN and the function node).

The value was calculated in the first case as 6.666/35.019 (~0.19) and in the second case 15.991/84.009 ~0.19). Thus, the final results are about five times more probable concerning defence over prosecution proposition.

DISCUSSION

Activity level assessments, standard formulae and associated BN structures do not exist within most current casework evaluations, which further challenges the custom case BN construction.

What should be required is that all statisticians preserve *justified* BN constructions. In other words – a thorough inspection of all structural component of a BN. Practically, this sort of system review can be strengthened by built-in functions of BN software that indicate graphical illustrations of dependencies and independencies among BN variables (Taylor et al, 2018b).

As we have concluded, there are no true or false probabilities, it gets pertinent to question about how particular probability values are allocated, and regarding which bases – so that an expert review can be conducted and open conversations about likelihood ratio assessment are welcomed.

We have to point out the necessity in observing the limitations of the information which are utilized to assign probabilities, the conditions under which the information was created.

The method of construction is adaptable enough that numerous cases with various circumstances can be assessed along these lines, yet standard enough that a statistician considering the BN (or beginning to consider how to develop a BN for a case) ought to advance quickly through architecture and comprehend the flow of information.

The statistician must remain in full control of the construction and application of the BN, as BNs ought to be built on a case per case basis to guarantee coherence and correct observation of the case features. BNs are expert support system tools, not proposed to substitute experts but to help them in their critical thinking.

We purposely picked the above-mentioned example as it features the significance of assessing cases considering action level proposition, instead of simply source level propositions. Regardless of whether the report proceeded to detail the potential reasons for false positives (or coincidental matches) for the findings, on the off chance that it did not put those in the framework of activities, one can perceive how the results would appear to a judge to firmly support the prosecution version of events over the defence version. Such a distinction can extremely affect the outcome of a trial.

CONCLUSION

When the source of the DNA is evident, biological results should be evaluated concerning activity level propositions. The premise of the estimation of the value of these outcomes must be stated in the report and justified. Undeniably, probabilities used to define likelihood ratios are dependent on data, knowledge and case information. Stating the limitations regarding case information and data to conclude probabilities is crucial. There are numerous obscure factors related to casework, as there are in the criminal trial or life as a whole. An admonition in the statement will strengthen the point that assessment is firmly conditioned upon the case-information that has been stated.

The utilization of Bayesian Networks as an exploratory instrument can detect data on how variables can impact on the value of the evidence. On the off chance that there is an absence of data that prevents evaluating the value of the results concerning activity level propositions, and on the off chance that transfer and background levels strongly impact on the case, then the court must be informed about this results (Gill et al., 2020).

In regards to the fairness and essentially to guarantee that the results have been evaluated in a robust way, there needs to be open access to data for both prosecution and defence scientists.

Furthermore, we believe the forensic science community will proceed with the LR as one potential, not normative or ideal tool in their communication with court members. We expect such perspectives will expand the priority concerning the development tools for presentations that follow the strict guidelines of scientific validity by concentrating on experimental or reproducible results, whilst helping the court members with an interpretation of the evidence (Lund & Iyer, 2017).

We have presented an example for BN construction that permits flexibility across cases, yet enough commonality and structure that the information flow in the BN is promptly perceived. Even though we have allocated this particular method, we need to stress that it is by no means the only manner in which BNs can be constructed. We share further the approach with the goal that others can benefit, and expand on what it has been presented (Taylor et al., 2018b).

One can argue that the proposed Bayesian Network construction is seemingly more detailed than necessary, and possibly influencing on the efficiency of the BN, stepping from the standard BNs construction, by utilizing them not similarly as methods for allocating and computing probabilities, but as graphical methods for explaining the assessment to other people. The utilization of such network structures makes it plausible, initially, to perceive what considerations have been made in the assessment of results.

REFERENCES

1. Biedermann, A., Taroni F. (2012) Bayesian networks for evaluating forensic DNA profiling evidence: a review and guide to literature. *Forensic Sci. Int. Genet.*, 6(2), 147-157.
2. Biedermann, A., C. Champod, Jackson, G., Gill, P., Taylor, D., Butler, J., Morling, N., Hicks, T., Vuille, J., Taroni, F. (2016) Evaluation of forensic DNA traces when propositions of interest relate to activities: analysis and discussion of recurrent concerns. *Front. Genet.*, 7, 215-220.
3. Buckleton, J., Bright, J.A., Taylor, D., Evett, I., Hicks, T., Jackson, G., Curran, J.M. (2014) Helping formulate propositions in forensic DNA analysis, *Sci. Justic*e, 54(4), 258-261.
4. ENFSI (2015) ENFSI Guideline for Evaluative Reporting in Forensic Science: Strengthening the Evaluation of Forensic Results across Europe (STEOFRAE)
5. https://www.unil.ch/esc/files/live/sites/esc/files/Fichiers%202015/ENFSI%20Guideline%20Evaluative%20Reporting
6. Evett, I.W., Jackson, G., Lambert J.A. (2000). More on the hierarchy of propositions: exploring the distinction between explanations and propositions. *Sci. Justice*, 40(1), 3-10.
7. Gill, P., Hicks, T., Butler, J. M., Connolly, E., Gusmão, L., Kokshoorn, B., Morling, N.,Van Oorschot, R., Parson, W., Prinz, M., Schneider, P. M., Sijen, T., Taylor, D. (2020). DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence - Guidelines highlighting the importance of propositions. Part II: Evaluation of biological traces considering activity level propositions. Forensic Science International: Genetics, 44, 102186.
8. Lund, S.P., Iyer, H. (2017) Likelihood Ratio as Weight of Forensic Evidence: A Closer Look. *Journal of Research of National Institute of Standards and Technology*, 22(27), 2017.
9. Meuwly, D., Ramos, D., Haraksim, R. (2017). A guideline for the validation of likelihood ratio methods used for forensic evidence evaluation. *Forensic Science International*, 276, 142-153.
10. National Resource Council (NRC) (1996a). Statistical Issues. In: N. Grossblatt (Ed.), *The Evaluation of Forensic DNA Evidence.* (2nd ed., pp. 125-165), Washington DC, USA: National Academy Press.
11. National Resource Council (NRC) (1996b). DNA Evidence in the Legal System. In: N. Grossblatt (Ed.), *The Evaluation of Forensic DNA Evidence.*. (2nd ed., pp. 166-211), Washington DC, USA: National Academy Press.
12. Neumann, C. & Ausdemore, M. (2020). Defence against the modern arts: the curse of statistics—Part II, 'Score-based likelihood ratios'. Law, Probability and Risk, 19(1), 21-42.
13. Pierson S. and Kafadar K. (2016). Statisticians and Forensic Science: A Perfect Match, *CHANCE*, 29(1), 4-8.

14. Stern, H. (2017). Statistical Issues in Forensic Science. *Annual Review of Statistics and Its Application*, 4(1), 225-244.
15. Taylor, D., Hicks, T., Champod, C. (2016) Using sensitivity analyses in Bayesian Networks to highlight the impact of data paucity and direct future analyses: a contribution to the debate on measuring and reporting the precision of likelihood ratios. *Sci. Justice*, 56(5), 402-410.
16. Taylor, D., Kokshoorn, B., Biedermann, A. (2018a) Evaluation of forensic genetics findings given activity level propositions: a review. *Forensic Sci. Int. Genet.,* 36, 34-49.
17. Taylor, D., Biedermann, A., Hicks, T., Champod, C. (2018b). A template for constructing Bayesian networks in forensic biology cases when considering activity level propositions. Forensic Science International: Genetics, 33, 136-146.
18. Weir, B. (2007). The rarity of DNA profiles. Annals of Applied Statistics, 1(2): 358–370.